

ABSTRACT

Input data queries directed at a plurality of target databases and originating from any of a plurality of sources are first converted to validated canonical forms, which are then used to query the target databases. Specifically, upon receiving an input data query, a relatively accurate reference database is selected based on the type of the input data. This reference is then queried for the input data with the intent of finding an exact matching record or a near-matching record that can be considered an exact match and thereby validating the input data. Otherwise, the requesting source is instructed to provide a new query. Once having a validated record, it is converted to a canonical form, which is then used to query the target databases intended to be searched. In a further embodiment, multiple reference databases are queried to determine a canonical form of the data or to determine multiple canonical forms of the data.